Random Assignment of Schools to Groups in the Drug Resistance Strategies Rural Project: Some New Methodological Twists

John W. Graham [A,B, C] **Jonathan Pettigrew** [A,D], **Michael L. Hecht** [A,C, D], **Michelle Miller-Day** [A,D], & **Janice Krieger** [E], [A] Penn State University,
[B] Dept of Biobehavioral Health, [C] Prevention Research Center, [D] Dept of Communication Arts & Sciences, [E] School of Communication, Ohio State University

Abstract (ID#: 1008172)
Random assignment to groups is the foundation for scientifically rigorous clinical trials. The theory behind random assignment depends on the "long run". This is especially key in group randomized trials in which whole schools are assigned to conditions, where it is rare that more than a few schools can be assigned to each condition. In the DRSR project, we assigned 39 rural PA and OH schools to 3 conditions (Rural, Classic, Control). Even with 13 schools per condition, it was still essential to maximize the chance that random assignment would achieve pretest equivalence on important study variables. We were able to collect data on several important school-level variables: a rurality index, number of grades in the school, enrollment per grade, percent white, percent receiving free/assisted lunch, and test scores. However, it is usually not possible to incorporate school-level drug use data into decisions about random assignment because access to the subject population is typically limited until after random assignment is complete. Fortunately, the PA Youth Study (PAYS) had recently been conducted, and drug use data from a random sample of students were available for 15 of the 26 PA schools involved in our study. Our goal was to follow a procedure outlined by Graham et al. (1984) to arrive at a single composite stratifying variable for randomization based on the school-level variables available. With drug use data for some of the schools, it was possible to use imputation (e.g., Schafer & Graham, 2002) to handle the incomplete data. In order to increase the number of schools, and the stability of the imputation analysis, we selected a sample of 28 additional PA schools (a) that met the rurality criteria for inclusion in our study, and (b) for which all school-level data, including PAYS drug use data, were available. The seven school-level variables were subjected to principal components analysis, and yielded two factors. These factors were averaged to form the Composite Inflate-Suppress (CIS) score which was the basis of stratification. The CIS score was broken into three strata within each state, and schools were assignment at random to the three program conditions (rural, classic, control) from within each stratum, within each state. As expected, program group membership was unrelated to the CIS score, the two factors making up the CIS score, and the seven items making up the factors. More importantly, program group membership was not significantly related to pretest measures of any of the main DVs (use of alcohol, cigarettes, marijuana, chewing tobacco; smallest p>.15), thus verifying that pretest equivalence was achieved.

Assignment to Groups: The Problem

Random assignment to groups

- Foundation of program evaluation, but depends on the "long run"
  - If enough units are assigned, program and control groups will be equivalent on all measured and unmeasured variables.
  - But long run hard to get with school-based studies.
  - Number of schools available for each group is typically small.
  - Pretest equivalence of groups is a concern.

Solutions

- Matching: Possible, but randomization is lost, and pretest equivalence is guaranteed only on variables involved in the matching.
- Stratified Randomization: Useful, but with small numbers, often only one stratification variable is feasible
- Graham et al. (1984) assignment procedure:
  - Take many school-level variables into account
  - Combine them to form a single stratifying variable

The Drug Resistance Strategies-Rural (DRSR) Program

The DRSR project studied adaptation processes for the school-based substance abuse prevention program known as *keepin' it REAL* (*kiR*; Hecht et al., 2003; 2007). We studied two approaches to adapting the *kiR* program.

- *Implementer adaptation*: We examined how teachers in largely Caucasian, rural schools adapted the original "**Classic**" version, which made use of the print and video materials originally developed for a largely Latino, urban population.
- *Developer Adaptation:* The developers adapted the curriculum resulting in the new, "**Rural**" version. New written and video materials involving rural PA students were developed so as to be more relevant for this more rural population.

Archival School-Level Data

- National Center for Educational Statistics,
  - Rurality index *(Xrurality)*.
  - Number of grades in the school *(Numgrades)*
  - Enrollment per grade *(Npergrade)*
  - Percent white *(Pctwhite)*
  - Percent receiving free or assisted lunch *(Pctlunch)*
- Pennsylvania Department of Education, Ohio Department of Education
  - Test scores *(Scores)*
- With the exception of the rurality index, these school-level variables were used by Graham et al. (1984), by Dent et al. (1993), or by both.

School-Level Drug Use Data

- Pennsylvania Survey of Youth
  - Included 15 of 26 PA schools in DRSR project

Missing Data Imputation

- Impute single data set from EM parameters (Graham et al., 2003)
  - We need parameter estimation, not hypothesis tests
  - Factor loadings, factor scores
- Help with imputation
  - Add 28 comparable schools with complete data

Drug Use Data

- Residual of Regression (Rdrugs):
  School Configuration (5-8 vs K-8) ➔ Drug use

The Assignment Procedure

Goal: Achieve Pretest Equivalence
Use Principal Components

- Data reduction
- Create single composite variable for stratifying

Table 2
Promax Rotated Factor Loadings

	Factor1 "rurality"	Factor2 "drugs"
numgrades	**.85**	-.02
pctwhite	**.75**	-.19
npergrade	**-.71**	-.25
xrurality	**.67**	.06
rdrugs (resid)	-.18	**.83**
scores	-.11	**-.78**
pctlunch	.08	**.68**
Factor Correlation		
Factor1	**1.00**	
Factor2	**.22**	**1.00**

Notes

- Annual Meeting of the Society for Prevention Research, Washington, DC, June 2, 2011.
- This research was supported by a grant from the National Institute on Drug Abuse, NIH Manuscript # NIHMS272843.
- Address: John W. Graham, Department of Biobehavioral Health, E-315 HHD Bldg, Penn State University, University Park, PA 16802.
- email: jgraham@psu.edu.

Composite Inflate-Suppress (CIS) Continuum

- IF high drugs in program, low drugs in control
  - SUPRESS: Program looks weaker than it really is
- IF high rural in program, low rural in control
  - Unclear what effect on interpretation of program effects
- But positive correlation between factors (r = .22) suggests averaging factors makes sense

CIS Score: Average of the two factors

Stratified Random Assignment of Schools to Groups (Numbers of Schools)

	Three Strata	***Pennsylvania Condition***			***Ohio Condition***		
CIS		***Cont.***	***Class***	***Rural***	***Cont.***	***Class***	***Rural***
• 0	Highest test scores	3	2	3	1	1	1
• 1	Medium test scores	3	3	3	2	2	2
• 2	Lowest test scores	3	3	3	2	2	2

Significance Tests for Group Equivalence

CIS-Related Variables

- No Significant Differences
  - CIS score,
  - Factors making up CIS score
  - Individual School-level variables making up factors

Pre-Test Measures of Main Drug Use Dependent Variables

- Used random intercepts model (Proc Mixed; controlled for ICC)
- No Significant Group Differences (smallest p = .15)
  - Lifetime Use of Alcohol, Cigarettes, Marijuana, Chewing Tobacco
  - Use in last 30 days of Alcohol, Cigarettes, Marijuana, Chewing Tobacco

Discussion

- Preliminary equivalence achieved
- 6 schools dropped out after assignment.
  - Replacing schools a simple matter of finding new schools in same CIS strata as schools that dropped out
- Pretest equivalence was assessed after replacement (so it worked)
- Missing data procedures offer interesting and useful extension of assignment procedure

References

Collins, Dent, C. W., Sussman, S., & Flay, B. R. (1993). The use of archival data to select and assign schools in a drug prevention trial. *Evaluation Review*, *17*, 159-181.
Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka & W. F. Velicer (Eds.). *Research Methods in Psychology* (pp. 87-114). Volume 2 of *Handbook of Psychology* (I. B. Weiner, Editor-in-Chief). New York: John Wiley & Sons.
Graham, J. W., Flay, B. R., Johnson, C. A., Hansen, W. B., and Collins, L. M. (1984). Group comparability: a multiattribute utility measurement approach to the use of random assignment with small numbers of aggregated units. *Evaluation Review*, *8*, 247-260.
Hecht, M. L., Marsiglia, F. F., Elek, E., Wagstaff, D. A., Kulis, S., Dustman, P., and Miller-Day, M. (2003). Culturally grounded substance abuse prevention: an evaluation of the keepin' it R.E.A.L. curriculum. Prevention Science, 4, 233-248.
Hecht, M.L. & Miller-Day, M. (2007). The drug resistance strategies project as translational research. Journal of Applied Communication Research, 35, 343-349.
Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, *7*, 147-177.